

#2
Jellison
5/15

501.39813X00

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant(s): T. YAZAKI ET AL.
Serial No.: Not Yet Assigned
Filed: March 20, 2001
For: PACKET SHAPER



CLAIM FOR PRIORITY

Assistant Commissioner for Patents
Washington, D.C. 20231

March 20, 2001

Sir:

Under the provisions of 35 U.S.C. and 37 CFR 1.55, the applicants hereby claim the right of priority based on:

Japan 2000-149811, filed May 17, 2000

The certified copy of said Japanese application is attached hereto.

Respectfully submitted,

ANTONELLI, TERRY, STOUT & KRAUS

Carl I. Brundidge
Registration No. 29,621

CIB:alw
(703) 312-6600

日 本 国 特 許 庁

PATENT OFFICE
JAPANESE GOVERNMENT

J1017 U.S. PTO
09/811504
03/20/01

別紙添付の書類に記載されている事項は下記の出願書類に記載されて
る事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed
in this Office.

出 願 年 月 日
Date of Application:

2000年 5月17日

出 願 番 号
Application Number:

特願2000-149811

出 願 人
Applicant(s):

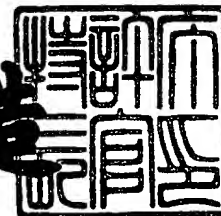
株式会社日立製作所

CERTIFIED COPY OF
PRIORITY DOCUMENT

2000年10月 6日

特許庁長官
Commissioner,
Patent Office

及 川 耕 造



出証番号 出証特2000-3081682

【書類名】 特許願

【整理番号】 H00009721A

【提出日】 平成12年 5月17日

【あて先】 特許庁長官 殿

【国際特許分類】 H04L 12/28

【発明者】

 【住所又は居所】 東京都国分寺市東恋ヶ窪一丁目 2 8 0 番地
株式会社日立製作所中央研究所内

 【氏名】 矢崎 武己

【発明者】

 【住所又は居所】 東京都千代田区神田駿河台四丁目 6 番地
株式会社日立製作所内

 【氏名】 相本 毅

【特許出願人】

 【識別番号】 000005108

 【氏名又は名称】 株式会社日立製作所

【代理人】

 【識別番号】 100075096

 【弁理士】

 【氏名又は名称】 作田 康夫

 【電話番号】 03-3212-1111

【手数料の表示】

 【予納台帳番号】 013088

 【納付金額】 21,000円

【提出物件の目録】

 【物件名】 明細書 1

 【物件名】 図面 1

 【物件名】 要約書 1

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 シェーピング装置

【特許請求の範囲】

【請求項 1】

パケットを蓄積するバッファと、前記パケットのキューイング優先度を判定する優先度判定手段と、前記キューイング優先度に基づいて前記バッファへのパケット蓄積・廃棄を判定する廃棄制御手段とを備え、

前記廃棄制御手段は、前記バッファへの非優先パケット蓄積・廃棄を判定する非優先パケットしきい値が設定されることを特徴とするシェーピング装置。

【請求項 2】

前記廃棄制御手段は、前記バッファへの優先パケット蓄積・廃棄を判定する優先パケットしきい値が設定されることを特徴とする請求項 1 記載のシェーピング装置。

【請求項 3】

前記バッファに蓄積しているパケット数をカウントするカウンタを備えることを特徴とする請求項 1 又は 2 記載のシェーピング装置。

【請求項 4】

複数のユーザからのパケットに対してそれぞれのユーザに対して保証している帯域分のパケットを優先パケットとする帯域監視部を有することを特徴とする請求項 1 又は 2 記載のシェーピング装置。

【請求項 5】

最低帯域を保証してパケットを送信するシェーピング装置において、パケットを蓄積する 1 つのバッファと、前記バッファに蓄積した順番でパケットを出力する出力制御手段と、前記バッファに蓄積する際のキューイング優先度を判定する優先度判定手段と、前記キューイング優先度に基づいて前記バッファへのパケット蓄積・廃棄を判定する廃棄制御手段を備えることを特徴とするシェーピング装置。

【請求項 6】

前記廃棄制御手段がキューイング優先度が「優先」のパケットを蓄積または廃

棄するかを判定する際に参照する前記バッファの最大値(閾値)を保存する優先パケット用閾値蓄積手段と、キューイング優先度が「非優先」のパケットを蓄積するか廃棄するかを判定する際に参照する前記バッファの最大値(閾値)を保存する非優先パケット用閾値蓄積手段と、前記バッファに蓄積しているパケット数をカウントするカウンタを備えることを特徴とする請求項5記載のシェーピング装置。

【請求項7】

前記優先度判定手段がパケットヘッダ内の情報等からパケットの一連のフローを検出して前記フローの識別子であるフロー識別子を判定するフロー検出手段と、帯域を監視するための制御情報である帯域監視制御情報から構成されるエントリを一つあるいは複数所持する帯域監視テーブルと、前記フロー識別子に対応するエントリを帯域監視テーブルから読み出す帯域監視テーブル制御手段と、前記帯域監視テーブル制御手段が読み出したエントリ内の帯域監視制御情報と現時刻を示すタイマーの値とパケットヘッダ内のパケット長に基づいてキューイング優先度を判定する監視結果判定手段を備えることを特徴とする請求項5乃至6記載のシェーピング装置。

【請求項8】

前記フロー検出手段が前記フロー識別子に加えて前記フロー内のパケットの重要度を判定し、前記監視結果判定手段が前記帯域監視制御情報と現時刻を示すタイマーの値とパケットヘッダ内のパケット長に加えて前記重要度に基づいてキューイング優先度を判定することを特徴とする請求項7記載のシェーピング装置。

【請求項9】

請求項5乃至6記載のシェーピング装置であって、前記優先度判定手段がパケットヘッダ内の情報等からパケットの一連のフローを検出して前記フローの識別子であるフロー識別子を判定するフロー検出手段と、帯域を監視するための制御情報であるパケットの深さ情報と帯域情報と前パケットの到着時刻情報と前パケット到着時のパケットの蓄積量情報から構成されるエントリを一つあるいは複数所持する帯域監視テーブルと、前記フロー識別子に対応するエントリを帯域監視テーブルから読み出す帯域監視テーブル制御手段と、読み出した帯域情報と前パケ

ットの到着時刻情報と前パケットの到着時のパケットの蓄積量情報と現時刻を示すタイマーの値から帯域監視直前のパケットの蓄積量を計算するパケット蓄積量判定手段と、前記帯域監視直前のパケットの蓄積量とパケットヘッダ内のパケット長とパケットの深さ情報に基づいてキューイング優先度を判定する監視結果判定手段を備えることを特徴とするシェーピング装置。

【請求項 1 0】

請求項 9 記載のシェーピング装置であって、前記フロー検出手段が前記フロー識別子に加えて前記フロー内のパケットの重要度を判定し、前記帯域監視テーブルが前記パケットの深さ情報として複数のパケットの深さを所持し、前記監視結果判定手段が前記帯域監視直前のパケットの蓄積量とパケットヘッダ内のパケット長とパケットの深さ情報に加えて前記重要度に基づいてキューイング優先度を判定することを特徴とするシェーピング装置。

【請求項 1 1】

パケットを蓄積するパケット蓄積FIFOと、前記パケットのキューイング優先度を決定する優先度決定手段と、前記キューイング優先度に基づいて前記パケット蓄積FIFOへのパケット蓄積・廃棄を判定する廃棄制御部とを備え、

前記廃棄制御部は、前記パケット蓄積FIFOへの非優先パケットの蓄積・廃棄を判定する際に参照する非優先パケット閾値が設定されることを特徴とするシェーピング装置。

【請求項 1 2】

前記廃棄制御部は、前記パケット蓄積FIFOへの優先パケットの蓄積・廃棄を判定する際に参照する優先パケット閾値が設定されることを特徴とする請求項 1 1 記載のシェーピング装置。

【請求項 1 3】

前記パケット蓄積FIFOに蓄積しているパケット数をカウントするカウンタを備えることを特徴とする請求項 1 1 又は 1 2 記載のシェーピング装置。

【請求項 1 4】

複数のユーザのパケットに対してそれぞれのユーザに保証している最低帯域分のパケットを優先パケットとする優先度決定手段を有することを特徴とする請求

項1乃至3記載のシェーピング装置。

【請求項 1 5】

前記優先度決定手段が、あるユーザのパケットを前記最低帯域を超えて受信した場合に、前記ユーザが予め決定した重要パケットを優先的に優先パケットと決定することを特徴とする請求項 1 4記載のシェーピング装置。

【請求項 1 6】

前記複数のユーザのパケットが同一の前記パケット蓄積FIFOに蓄積されることを特徴とする請求項 1 1乃至 1 5記載のシェーピング装置。

【請求項 1 7】

最低帯域を保証してパケットを送信するシェーピング装置において、パケットを蓄積する1つのパケット蓄積FIFOと、前記パケット蓄積FIFOに蓄積した順番でパケットを出力する出力制御手段と、前記パケット蓄積FIFOに蓄積する際のキューイング優先度を決定する優先度決定手段と、前記キューイング優先度に基づいて前記パケット蓄積FIFOへのパケット蓄積・廃棄を判定する廃棄制御部を備えることを特徴とするシェーピング装置。

【請求項 1 8】

前記廃棄制御部がキューイング優先度が「優先」のパケットを蓄積または廃棄するかを判定する際に参照する前記パケット蓄積FIFOの最大値(閾値)を保存する優先パケット用閾値蓄積手段と、キューイング優先度が「非優先」のパケットを蓄積するか廃棄するかを判定する際に参照する前記パケット蓄積FIFOの最大値(閾値)を保存する非優先パケット用閾値蓄積手段と、前記パケット蓄積FIFOに蓄積しているパケット数をカウントするカウンタを備えることを特徴とする請求項 1 7記載のシェーピング装置。

【請求項 1 9】

前記優先度決定手段がパケットヘッダ内の情報等からパケットの一連のフローを検出して前記フローの識別子であるフロー識別子を判定するフロー検出手段と、帯域を監視するための制御情報である帯域監視制御情報から構成されるエントリを一つあるいは複数所持する帯域監視テーブルと、前記フロー識別子に対応するエントリを帯域監視テーブルから読み出す帯域監視テーブル制御手段と、前

記帯域監視テーブル制御手段が読み出したエントリ内の帯域監視制御情報と現時刻を示すタイマーの値とパケットヘッダ内のパケット長に基づいてキューイング優先度を決定する監視結果判定手段を備えることを特徴とする請求項17乃至18記載のシェーピング装置。

【請求項20】

前記フロー検出手段が前記フロー識別子に加えて前記フロー内のパケットの重要度を判定し、前記監視結果判定手段が前記帯域監視制御情報と現時刻を示すタイマーの値とパケットヘッダ内のパケット長に加えて前記重要度に基づいてキューイング優先度を決定することを特徴とする請求項19記載のシェーピング装置。

【請求項21】

請求項17乃至18記載のシェーピング装置であって、前記優先度決定手段がパケットヘッダ内の情報等からパケットの一連のフローを検出して前記フローの識別子であるフロー識別子を判定するフロー検出手段と、帯域を監視するための制御情報であるパケットの深さ情報と帯域情報と前パケットの到着時刻情報と前パケット到着時のパケットの蓄積量情報から構成されるエントリを一つあるいは複数所持する帯域監視テーブルと、前記フロー識別子に対応するエントリを帯域監視テーブルから読み出す帯域監視テーブル制御手段と、読み出した帯域情報と前パケットの到着時刻情報と前パケットの到着時のパケットの蓄積量情報と現時刻を示すタイマーの値から帯域監視直前のパケットの蓄積量を計算するパケット蓄積量判定手段と、前記帯域監視直前のパケットの蓄積量とパケットヘッダ内のパケット長とパケットの深さ情報に基づいてキューイング優先度を決定する監視結果判定手段を備えることを特徴とするシェーピング装置。

【請求項22】

請求項21記載のシェーピング装置であって、前記フロー検出手段が前記フロー識別子に加えて前記フロー内のパケットの重要度を判定し、前記帯域監視テーブルが前記パケットの深さ情報として複数のパケットの深さを所持し、前記監視結果判定手段が前記帯域監視直前のパケットの蓄積量とパケットヘッダ内のパケット長とパケットの深さ情報に加えて前記重要度に基づいてキューイング優先度を決定す

ることを特徴とするシェーピング装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、パケットの送信帯域を成形するシェーピング装置に関する。

【0002】

【従来の技術】

インターネットで用いられているパケット型通信方式では、多数のユーザからのパケットが同じ回線を共用使用できるため、帯域あたりのコストを低く抑えることが出来る。このインターネットにおいて、従来の電話網や企業網が実現していた低遅延時間や低廃棄率等の通信品質（QoS: Quality of Service）を実現する要求がでてきた。

【0003】

QoSを実現するサービスの一つである最低帯域保証サービスは、各ユーザ（契約者）毎に契約された最低帯域が保証され、この最低帯域を超過したパケットに関してはネットワーク資源が許す限り転送される。

【0004】

例えば、図2のインターネット200で最低帯域保証サービスが提供されており、エッジルータB204に直接接続する5つの端末（端末E235、端末F236、端末G237、端末H238、端末I239）が受信する最低帯域が、それぞれユーザE、ユーザF、ユーザG、ユーザH、ユーザIと契約されている場合について考える。エッジルータA201に直接接続する端末A231、端末B232、端末C233、端末D234の送信したトラヒックはエッジルータA201、バックボーンルータA202、バックボーンルータB203、エッジルータ204Bにより端末E235、端末F236、端末G237、端末H238、端末I239まで転送されている。バックボーンルータB203からエッジルータB204へ至る回線がボトルネックとなっている場合、バックボーンルータB203は前記回線用のシェーパを所持し、各ユーザ毎に最低帯域を保証して転送する。

【0005】

最低帯域保証するシェーパのスケジューリング方式としては例えばWFQ(Weight

ed Fair Queueing)と呼ばれるスケジューリング方式が知られている。WFQのアルゴリズムは、SCFQ(Self Clocked Fair Queueing)と呼ばれるアルゴリズムがS.Golestani, A Self-Clocked Fair Queueing Scheme for Broadband Applications, In proc. of INFOCOM94, pp.636-646,1994.に記述されている(従来技術1)。従来技術1ではシェーパは複数のキュー i ($=1\sim N$)を所持し、前記キュー i 毎に使用帯域に比例した重み W_i と変数 F_i を持ち、パケットがシェーパに到着すると変数 F_i の更新を行う。パケット出力時には、パケットが蓄積しているキュー i の先頭パケットの変数 F_i の中から最小の F_i を持つキュー i のパケットをサービスする。変数の F_i の更新は

$$F_i = L/W_i + \max(F_i, V(t_a))$$

に従って行われる。ここで、 L は到着パケットのパケット長、 t_a はパケットの到着時刻、 $V(t)$ は時刻 t におけるキュー i のパケット変数 F_i の値を返す関数である。この方式では、キュー i からは $W_i / (W_i \text{の総和}) \times \text{回線帯域}$ 分の最低帯域が保証される。

【 0 0 0 6 】

【発明が解決しようとする課題】

従来技術1のシェーパには2つの問題点がある

第一の問題点は、ユーザ数が増加すると高速にシェーピングできない点である。従来技術1のシェーパは各ユーザ(=契約者)のトラフィックをユーザ毎のキューにキューイングする。パケットを送信するキューを判定する時間はユーザ数が増加すると共に増加するため、高速にシェーピングできない。

【 0 0 0 7 】

第二の問題点は、ユーザが最低帯域を効率良く利用できない点である。前述した図2のインターネット200でこの問題点を説明する。端末A231、端末B232、端末C233、端末D234からユーザEの端末E235行きのトラフィックが図9の様にバックボーンルータB203に入力する。また、バックボーンルータB203のシェーパはユーザEの最低帯域を保証してシェーピングしている。バックボーンルータB203において全ユーザのトラフィックの総和が図10に示す様に変化すると、前記シェーパが出力するトラフィックは図11の様に变化する。時間 t_a では全ユーザのトラフィックの総和が

回線帯域を下回っている為、ユーザEへのトラヒックはユーザ毎のキューに蓄積されると直ぐに出力される。また、時間tbでもユーザEの帯域が最低帯域を下回っているためユーザEのトラヒックは全て出力される。時間tcでは全ユーザのトラヒックの総和が回線帯域を超え、ユーザEの帯域が最低帯域を超過しているため、ユーザEのパケットはユーザ毎のキューに蓄積され、tcの期間が長いとキュー長が増加しパケットは廃棄されてしまう。シェーパはパケットの重要度を考慮していないため、ユーザが受信端末まで必ず到達させたい重要なパケットを廃棄し、それ以外のパケットを転送する場合もある。従来技術1のシェーパはユーザにとって重要なパケットを廃棄してそれ以外のパケットを送信してしまい、ユーザが最低帯域を効率良く利用出来ない場合がある。なお、本発明ではユーザが受信端末まで必ず到達させたい重要なパケットを重要パケットと、それ以外のパケットを非重要パケットと呼ぶ。

【0008】

以上に述べた様に従来の技術のシェーパは多くのユーザの最低帯域を保証して高速にシェーピングすることができなかった。

【0009】

本発明の第一の目的は、多くのユーザの最低帯域を保証して高速にシェーピングするシェーピング装置を提供することである。

【0010】

また、従来の技術のシェーパは、重要パケットを保護して、最低帯域を効率良く利用できないため、ユーザが最低帯域を効率良く利用できなかった。

【0011】

本発明の第二の目的は、重要パケットを保護して、最低帯域を効率良く利用するシェーピング装置を提供することである。

【0012】

【課題を解決するための手段】

上記第一の目的を達成するために、本発明のシェーピング装置は、パケットを蓄積するパケット蓄積FIFOと、前記パケットのキューイング優先度を決定する優先度決定手段と、前記キューイング優先度に基づいて前記パケット蓄積FIFOへの

パケット蓄積・廃棄を判定する廃棄制御部とを備え、複数のユーザのパケットが同一の前記パケット蓄積FIFOに蓄積されることを特徴とする。本発明のシェーピング装置は一つのパケット蓄積FIFO520からパケットを出力するため、複数のバッファからパケットを出力するバッファを選択する必要がなく高速にシェーピングを行う事ができる。

【 0 0 1 3 】

上記第二の目的を達成するために、本発明のシェーピング装置は、複数のユーザのパケットに対してそれぞれのユーザに保証している最低帯域分のパケットを優先パケットとする優先度決定手段を備え、前記優先度決定手段があるユーザのパケットを前記最低帯域を超えて受信した場合に、前記ユーザが予め決定した重要パケットを優先的に優先パケットと決定することを特徴とする。本発明のシェーピング装置は重要パケットを保護して転送するため、最低帯域を効率良く利用することができる。

【 0 0 1 4 】

その他の本発明が解決しようとする課題、その解決手段は、本発明の「発明の実施の形態」の欄及び図面で明らかにされる。

【 0 0 1 5 】

【発明の実施の形態】

まず、本発明のシェーピング機能を所持するルータの概要動作を図1乃至図4を用いて説明する。図1は本発明を適用したルータ100を示す。ルータ100はパケットが入力するN個の入力回線110-i (i=1~N)とパケットの受信処理を行うパケット受信回路120-iと、ルーティング処理部130と、パケットをスイッチングするパケット中継処理手段140と、出力回線毎のシェーパ部500-j (j=1~N)と送信処理を行うパケット送信回路150-jと、パケットが出力されるN個の出力回線160-jより構成される。

【 0 0 1 6 】

図3はインターネット200におけるパケットのフォーマットの一例を示す。インターネット200におけるパケットはヘッダ部310とデータ部320から構成される。ヘッダ部310は送信元アドレス（送信端末のアドレス）である送信元IPアドレス（S

source IP Address: 以下「SIP」という。) 311と、宛先アドレス(受信端末のアドレス)である宛先IPアドレス(Destination IP Address: 以下「DIP」という。) 312と、送信元のプロトコル(=上位アプリケーション)を表す送信元ポート(Source Port: 以下「SPORT」という。) 313と宛先のプロトコルを表す宛先ポート(Destination Port: 以下「DPORT」という。) 314から構成される。また、データ部320はユーザデータ321から構成される。ヘッダ部310には前記情報以外にサービス識別子(TOS: Type of Service)等の情報も格納されているが、前記情報と同様に後述の処理を実行することができる。

【 0 0 1 7 】

図4は本発明を適用したルータ100内部におけるパケットのフォーマットを示す。ルータ100内部におけるパケットのフォーマットはインターネット200におけるパケットのフォーマットに内部ヘッダ部330が備わる。この内部ヘッダ部330はパケットのバイト長を表すパケット長331とパケットが入力した回線の識別子である入力回線番号332と、パケットが出力される回線の識別子である出力回線番号333から構成される。

【 0 0 1 8 】

本発明が想定しているネットワーク構成を図2に示した。図2のインターネット200では最低帯域保証サービスが提供されており、エッジルータB204に直接接続する5つの端末(端末E235、端末F236、端末G237、端末H238、端末I239)が受信する最低帯域がそれぞれユーザE、ユーザF、ユーザG、ユーザH、ユーザIと契約されている。端末A231、端末B232、端末C233、端末D234がパケットを送信し、端末E235、端末F236、端末G237、端末H238、端末I239が受信している。バックボーンルータB203からエッジルータB204へ至る回線の帯域は各ユーザと契約している最低帯域の総和よりも大きいがボトルネックとなっている。本発明を適用したルータ100がバックボーンルータB203として使用され、エッジルータB204へ至る出力回線160-jのシェーパ部500-jが各ユーザへの最低帯域を保証するようにシェーピングを実行している。

【 0 0 1 9 】

本発明を適用したルータ100の入力回線110-iよりパケットが入力するとパケッ

ト受信回路120-iは内部ヘッダ部330を付加し、前記パケットのバイト長をカウントしてパケット長331(単位はByte)に書き込む。さらに、パケットが入力した入力回線110-iの回線番号iを入力回線番号332に書き込み、ルーティング処理部130へパケットを送信する。なお、この時点では出力回線番号333は無意味な値となっている。ルーティング処理部130はパケットを受信するとDIP312よりパケットを出力する出力回線160-jを判定し、出力回線160-jの回線番号jを出力回線番号333に書き込みパケット中継処理手段140に送信する。パケット中継処理手段140は出力回線番号333に従いパケットをスイッチングし、出力回線毎のシェーパ部500-jに送信する。シェーパ部500-jはユーザ毎の最低帯域を保証して、パケットをパケット送信回路150-jに送信する。パケット送信回路150-jはパケットを受信すると、内部ヘッダ部330を削除し、出力回線160-jにパケットを送信する。

【 0 0 2 0 】

次に、本発明のシェーパ部500の詳細動作について説明する。図5に本発明のシェーパ部500のブロック図を示した。シェーパ部500はパケットを蓄積するパケット蓄積FIFO520と、前記パケット蓄積FIFO520に蓄積した順番でパケットを出力するパケット出力制御部530と、前記パケット蓄積FIFO520へのパケット入力や廃棄を制御する廃棄制御部510と、前記パケット蓄積FIFO520に蓄積する時のキューイング優先度を決定する帯域監視部600より構成される。本発明では帯域監視部600がキューイング優先度を「優先」と決定するパケットを優先パケットと、「非優先」と決定するパケットを非優先パケットと呼ぶ。優先パケットは輻輳が発生してパケット蓄積FIFO520にパケットが大量に蓄積された場合に、前記パケット蓄積FIFO520に優先的に蓄積される。

【 0 0 2 1 】

廃棄制御部510はパケットを一旦蓄積する一時蓄積バッファ511と、パケット蓄積FIFO520に蓄積しているパケット数をカウントするFIFOカウンタ513と、前記FIFOカウンタ513の優先パケット用のパケット蓄積FIFO520の閾値を格納する優先パケット用閾値蓄積手段514と、非優先パケット用のパケット蓄積FIFO520の閾値を格納する非優先パケット用閾値蓄積手段515と、廃棄判定回路512より構成される。なお、優先パケット用閾値蓄積手段514に蓄積される閾値は非優先パケッ

ト用閾値蓄積手段515に蓄積される閾値よりも大きい。

【 0 0 2 2 】

パケット中継処理手段140よりパケットが入力すると、ヘッダ部310、データ部320、内部ヘッダ部330の全てが一時蓄積バッファ511に、ヘッダ部310、内部ヘッダ部330が帯域監視部600に蓄積される。帯域監視部600は蓄積した情報を基にキューイング優先度を決定する。ユーザ毎に入力するパケットの帯域を監視し、最低帯域以内のパケットに対して「優先」と判定し、超過したパケットに対して「非優先」と判定する。さらに、この判定情報から構成されるキューイング優先度12を廃棄判定回路512に送信する。なお、帯域監視部600の動作の詳細については後述する。

【 0 0 2 3 】

廃棄制御部510の廃棄判定回路512は前記キューイング優先度12とFIFOカウンタ513の値と優先パケット用閾値蓄積手段514内の閾値と非優先パケット用閾値蓄積手段515内の閾値から一時蓄積バッファ511内のパケットを廃棄するか否かを判定する。キューイング優先度12が「優先」の場合、
優先パケット用閾値蓄積手段514内の閾値 $>$ FIFOカウンタ513の値
の時に「蓄積」と判定して前記パケットをパケット蓄積FIFO520へ送信するように一時蓄積バッファ511に指示する。一時蓄積バッファ511はこの信号を受信するとパケット蓄積FIFO520にパケットを送信し、パケット蓄積FIFO520はこのパケットを蓄積する。

【 0 0 2 4 】

優先パケット用閾値蓄積手段514内の閾値 $=$ FIFOカウンタ513の値
の時には「廃棄」と判定して、一時蓄積バッファ511に前記パケット送信の指示を行わない。一時蓄積バッファ511は「廃棄」と判定したパケットの情報に次に到着したパケットの情報を上書きする。一方、キューイング優先度12が「非優先」の場合、

非優先パケット用閾値蓄積手段515内の閾値 $>$ FIFOカウンタ513の値
の時に「蓄積」と判定し、

非優先パケット用閾値蓄積手段515内の閾値 \leq FIFOカウンタ513の値

の時には「廃棄」と判定する。廃棄判定回路512は判定結果(「蓄積」か「廃棄」か)を廃棄判定結果13としてFIFOカウンタ513に送信する。FIFOカウンタ513は廃棄判定結果13が「蓄積」の場合には、蓄積情報に1加算し、「廃棄」の場合には加算を行わない。

【 0 0 2 5 】

パケット出力制御部530はパケット蓄積FIFO520に蓄積した順番に回線帯域でパケットを送信するようにパケット送信起動信号14を送信し、この信号を受信したパケット蓄積FIFO520はパケット送信回路150へパケットを送信する。また、FIFOカウンタ513は前記パケット送信起動信号14を受信すると蓄積情報から1減算する。本発明では、パケット出力制御部530は回線帯域でパケットを送信するようにパケット送信起動信号14を送信するが、ネットワーク運用者によって設定された回線帯域以下の帯域(例えば、回線帯域の半分)で送信しても良い。

【 0 0 2 6 】

本発明のシェーパ部500-jは一つのパケット蓄積FIFO520からパケットを出力するため、複数のバッファからパケットを出力するバッファを選択する必要がなく高速にシェーピングを行う事ができる。

【 0 0 2 7 】

ユーザE(端末E235)行きのトラヒックが図9に示す様にシェーパ部500-jに入力し、全ユーザのトラヒックの総和が図10に示す様にシェーパ部500-jに入力する場合、出力回線160-jより出力されるユーザE行きのトラヒックは図11のようになる。時間taでは全ユーザのトラヒックの総和が回線帯域を下回っている為、ユーザEのパケットはパケット蓄積FIFO520に蓄積されて出力される。時間tb、tcでは全ユーザのトラヒックの総和が回線帯域を超えている為、パケット蓄積FIFO520に非優先パケット用閾値蓄積手段515内の値を超えてパケットが蓄積される。この時、非優先パケットは蓄積されず、優先パケットのみパケット蓄積FIFO520に蓄積される。パケット蓄積FIFO520からは回線帯域でパケットは出力され、各ユーザの優先パケットは最低帯域以下で入力する。各ユーザの最低帯域の総和は回線帯域以下であるから、優先パケットは廃棄されずにパケット蓄積FIFO520に蓄積され、出力回線160-jより出力される。時刻tbではユーザEの帯域は最低帯域以下

であるため、最終的にユーザEのパケットは出力回線160-jより出力される。時間tcではユーザEの最低帯域以下のパケットは優先パケットと判定され出力回線160-jより出力されるが、最低帯域以上のパケットは非優先パケットと判定されて廃棄される。

【 0 0 2 8 】

次に、帯域監視部600の詳細動作について説明する。時間tcでは帯域監視部600は重要パケットを選択的に優先パケットと判定する。この機能によりシェーパ部500は重要パケットを保護することが出来る。帯域監視のアルゴリズムとして固定長パケットであるセルの監視アルゴリズムであるcontinuousLeaky Bucket Algorithm(以下リーキーバケットアルゴリズム)を可変長パケットの帯域監視用に拡張したアルゴリズムを使用する。リーキーバケットアルゴリズムに関しては例えばThe ATM Forum Specification version 4.1の4.4.2章に記載されている。リーキーバケットアルゴリズムはある深さを持った穴の空いた漏れバケツのモデルで、バケツに水が入っている間は監視帯域で水は漏れ、セル到着時にはバケツに1セル分の水が注ぎ込まれる。セルの到着揺らぎを許容するためにバケツに深さを持ち、バケツが溢れない内は入力セルは遵守と、溢れると違反と判定される。本発明ではパケット到着時に注ぎ込む水の量を可変とすることにより、可変長パケットの帯域監視を実現する。

【 0 0 2 9 】

図6に帯域監視部600のブロック図を示す。帯域監視部600は帯域監視テーブル制御部650と、バケツ蓄積量判定部610と、監視結果判定部620と、帯域監視テーブル630と、フロー検出手段640より構成される。フロー検出手段640は帯域監視を行うために、入力パケット毎にヘッダ内の情報等によりフロー識別子を判定する。帯域監視部600は前記フロー識別子に対応する帯域監視制御情報を用いて帯域監視を実行する。なお、本発明では、ヘッダ内の情報等を組み合わせて作成したパケット識別の条件をフロー条件と、フロー条件に一致する一連のトラヒックをフローと、フロー条件に入力パケットが一致するか否かを判定してフローの識別を行うことをフロー検出と呼ぶ。

【 0 0 3 0 】

図7に帯域監視テーブル630のフォーマットを示す。前記帯域監視テーブル630はM個の帯域監視制御情報700-k(k=1~M)をより構成される。帯域監視部600は一つの前記帯域監視制御情報700-kにより1ユーザの帯域監視を実行する。帯域監視制御情報700-kは重要パケット用のパケツの深さTHR-A701-k(Byte) (Threshold-A)と、非重要パケット用のパケツの深さTHR-B702-k(Byte) (Threshold-B)と、パケツから水が漏れる速度であり監視レートを表すPOLR703-k(Byte/sec) (Policing Rate)と、同一の帯域監視制御情報700-k(k=1~M)を参照するパケツの水が前回パケツに蓄積した時刻TS704-k(sec) (Time Stamp)と、時刻TS704-kにパケツに蓄積していた水の量であるCNT705-k(Byte) (Count)より構成される。なお、パケツの深さを表すTHR-A701-kとTHR-B702-kは、 $THR-A701-k \geq THR-B702-k$ の関係がある。

【 0 0 3 1 】

図8に帯域監視部600のフローチャートを示す。帯域監視部600の処理は帯域監視開始処理800、パケツ蓄積量判定処理810、監視結果判定処理820である。後の2処理はそれぞれパケツ蓄積量判定部610と、監視結果判定部620が主に実行する。

【 0 0 3 2 】

帯域監視部600がヘッダ部310と内部ヘッダ部330から構成されるパケツヘッダ情報11を受信すると、監視結果判定部620のパケツ長蓄積手段622はパケツ長331を、フロー検出部640は入力回線番号332と、出力回線番号333と、SIP311と、DIP312と、SPORT313と、DPORT314を蓄積する(ステップ801)。ステップ802では、フロー検出部640は蓄積した情報に基づいてフロー検出を行う。ここでは、一時蓄積バッファ511内のパケツのフロー識別子と重要または非重要パケツかを判定する。さらに、前記フロー識別子から構成されるフロー識別子情報15を帯域監視テーブル制御部650の帯域監視テーブル制御回路651へ、パケツの重要度を表すパケツ重要度情報17を監視結果判定部620のパケツ重要度蓄積手段624へ送信する。

【 0 0 3 3 】

このフロー識別子情報15とユーザは1対1で対応する。図2のインターネット200ではフロー検出部640は、例えば、端末E235、端末F236、端末G237、端末H238、端末I239のIPアドレス(DIP312)によってユーザを識別してフロー識別子を判定し、TCPのポート番号(SPORT313、DPORT314)からアプリケーションを識別してパケットの重要度を判定する。

【 0 0 3 4 】

帯域監視テーブル制御回路651は前記フロー識別子情報15を受信すると、帯域監視テーブル630のアドレスを作成して帯域監視制御情報700-kを読み出し、THR-A701-kとTHR-B702-kを監視結果判定部620のTHR蓄積手段623に、POLR703-kとTS704-kとCNT705-kをバケツ蓄積量判定部610のそれぞれPOLR蓄積手段613、TS蓄積手段614、CNT蓄積手段615に蓄積する(ステップ803)。

【 0 0 3 5 】

バケツ蓄積量判定処理810では、バケツ蓄積量判定部610はパケット入力直前のバケツの水の量(バケツ蓄積量)を判定する。まず、バケツ蓄積量判定回路611は現時刻をカウントするタイマー612の値(単位はsec)とTS蓄積手段614内のTS704-k(sec)との差分を計算し、バケツに水が前回蓄積されてから経過した経過時間(sec)を計算する(ステップ811)。次に経過時間(sec)にPOLR蓄積手段613内のPOLR703-k(Byte/sec)を乗じて、バケツに水が前回蓄積されてから漏れた水の量(バケツ減少量)を計算する(ステップ812)。さらに、CNT蓄積手段615内のCNT705-kからバケツ減少量を減算してパケットが入力する直前のバケツ蓄積量を判定する(ステップ813)。前記バケツ蓄積量の正負を判定し(ステップ814)、判定結果が負の場合にはバケツ蓄積量を"0"(バケツは空)に修正する(ステップ815)。

【 0 0 3 6 】

監視結果判定処理820では、監視結果判定部620の監視結果判定回路621は一時蓄積バッファ511内のパケットに相当する水がバケツに入るか否かを判定してパケットの優先度を決定する。まず、監視結果判定回路621はバケツ蓄積量判定処理810で判定されたバケツ蓄積量(Byte)にパケット長(Byte)を加算する(ステップ821)。次に、パケット重要度蓄積手段624はフロー検出部640が送信しているパケ

ット重要度情報17を蓄積する。この蓄積情報に基づき帯域監視処理は分岐する(ステップ822)。監視結果判定回路621は前記蓄積情報が「重要」の場合にはTHR蓄積手段623に蓄積されている重要パケット用のバケツの深さTHR-A701-kと前記加算値との大小比較を行う(ステップ823)。バケツ蓄積量+パケット長>THR-A701-k、即ち、バケツが溢れてしまう時には、一時蓄積バッファ511に蓄積されているパケットを違反パケットと判定して「違反」を表す帯域監視結果情報18を帯域監視テーブル制御部650の帯域監視テーブル制御回路651に、「非優先」を表すキューイング優先度12を廃棄判定回路512に送信する(ステップ826)。一方、バケツ蓄積量+パケット長 \leq THR-A701-kの時には、入力パケットを遵守パケットと判定し、「遵守」を表す帯域監視結果情報18を帯域監視テーブル制御回路651に、「優先」を表すキューイング優先度12を廃棄制御部510の廃棄判定回路512に送信し、「バケツ蓄積量+パケット長」の値をバケツ蓄積量情報16として帯域監視テーブル制御回路651に送信する(ステップ825)。ステップ822の参照結果が「非重要」の場合にはバケツ蓄積量+パケット長の値がTHR蓄積手段623に蓄積されている非重要パケット用のバケツの深さTHR-B702-kと前記加算値との大小比較を行う(ステップ824)。バケツ蓄積量+パケット長 \leq THR-B702-kの時には前記ステップ825を、バケツ蓄積量+パケット長>THR-B702-kの時には前記ステップ826を実行する。

【 0 0 3 7 】

帯域監視テーブル制御回路651は「遵守」を表示した帯域監視結果情報18を受信すると、バケツ蓄積量情報16とタイマー612の値を、それぞれ帯域監視直後のバケツ蓄積量およびパケットの到着時刻として、帯域監視テーブル630のCNT705-kとTS704-kに書き込む(ステップ827)。帯域監視テーブル制御回路651は「違反」を表示した帯域監視結果情報18を受信すると前記ステップ827を行わない。以上の処理が終了すると帯域監視は終了する(ステップ828)。

【 0 0 3 8 】

ステップ822およびステップ824は本発明固有の処理である。帯域監視部600は2つのバケツの深さTHR-A701-k、THR-B702-kを所持し、ステップ822およびステップ824を実行することにより、重要パケットを選択的に優先パケットと判定する。前

記ユーザE行きのトラヒックの重要パケットと非重要パケットの内訳が図12の場合について考える。この場合、帯域監視部600の判定結果は図13のようになる。メッシュ部が優先パケット、白抜き部が非優先パケットである。時間tbでは最低帯域以下でパケットが入力される為、パケットに水は蓄積されず、全てのパケットが「優先」と判定される。一方、時間ta、tcでは最低帯域以上にパケットが入力される為、非重要パケット用のパケットの深さTHR-B702-kまで水が蓄積される。この時、非重要パケットは蓄積されず「非優先」と判定され、重要パケットのみがパケットに蓄積され「優先」と判定される。

【 0 0 3 9 】

出力回線160-jより出力されるトラヒックを図14に示した。時間taでは全ユーザのトラヒックの総和が回線帯域を下回っている為(図10参照)、ユーザE行きのパケットはパケット蓄積FIFO520に蓄積されると直ぐに出力される。時間tcでは全ユーザのトラヒックの総和が回線帯域を超えている為、パケット蓄積FIFO520に非優先パケット用閾値蓄積手段515内の閾値を超えてパケットが蓄積される。この時、非優先パケットはパケット蓄積FIFO520に蓄積されず、優先パケットのみ蓄積される。パケット蓄積FIFO520からは回線帯域でパケットは出力され、優先パケットは回線帯域以下で入力する為、優先パケットは廃棄されずにパケット蓄積FIFO520に蓄積されて最終的に出力回線160-jより出力される。このため、ユーザEの全ての重要パケットは出力回線160-jより出力される。時間tbではユーザE行きのパケットは全て優先パケットと判定されるため、全て出力回線160-jより出力される。

【 0 0 4 0 】

以上に説明した様にシェーパ部500の帯域監視部600が重要パケットを選択的に優先パケットと判定して廃棄制御部510が前記優先パケットをパケット蓄積FIFO520に優先的に蓄積することにより、シェーパ部500は重要パケットを保護して転送することができる。

【 0 0 4 1 】

これまで、ルータ100の動作について説明した。ATM交換機やフレームリレー交換機の場合、ヘッダ部310はVPI/VCIやDLCIといったコネクション識別子やCLPやDE

といった廃棄優先度情報から構成される。これに対応してルーティング処理部130はDIP312の代わりに前記コネクション識別子よりパケットを出力する出力回線160-jを判定し、フロー検出部640はSIP311と、DIP312と、SPORT313と、DPORT314の代わりにコネクション識別子や廃棄優先度情報を使用してフロー検出を行う。

【0042】

【発明の効果】

本発明のシェーパ部500を備えたルータ100は多くのユーザの最低帯域を保証して高速にシェーピングすることができる。ネットワーク運用者はこのルータ100を用いることにより、高速なネットワークにおいて多くのユーザに最低帯域保証サービスを提供することができる。

【0043】

さらに、本発明のシェーパ部500を備えたルータ100は重要パケットを保護して、最低帯域を効率良く利用するシェーピングを行うことができる。ネットワーク運用者はこのルータ100を用いることにより、ユーザにとって重要なパケットを保護して送信する「最低帯域を有効に利用できるサービス」を提供することができる。

【図面の簡単な説明】

【図1】

本発明を適用したルータ100の構成を示すブロック図。

【図2】

ネットワークの構成図。

【図3】

インターネット200におけるパケットのフォーマットを示す図。

【図4】

本発明を適用したルータ100におけるパケットのフォーマットを示す図。

【図5】

本発明のシェーパ部500の構成を示すブロック図。

【図6】

本発明の帯域監視部600の構成を示すブロック図。

【図 7】

帯域監視テーブル630のフォーマットを示す図。

【図 8】

本発明を適用した帯域監視部600のフローチャート。

【図 9】

シェーパ部500に入力するトラヒックで端末E235行きトラヒックの時間変化を表す図。

【図 1 0】

シェーパ部500に入力する全ユーザのトラヒックの総和の時間変化を表す図。

【図 1 1】

出力回線160-jから出力するトラヒックで端末E235行きトラヒックの時間変化を表す図。

【図 1 2】

シェーパ部500に入力するトラヒックで、端末E235行きの重要パケットと非重要パケットの内訳の時間変化を表す図。

【図 1 3】

帯域監視部600が「優先」または「非優先」と判定するトラヒックの時間変化を表す図。

【図 1 4】

出力回線160-jから出力するトラヒックで端末E235行きトラヒックの時間変化を表す図。

【符号の説明】

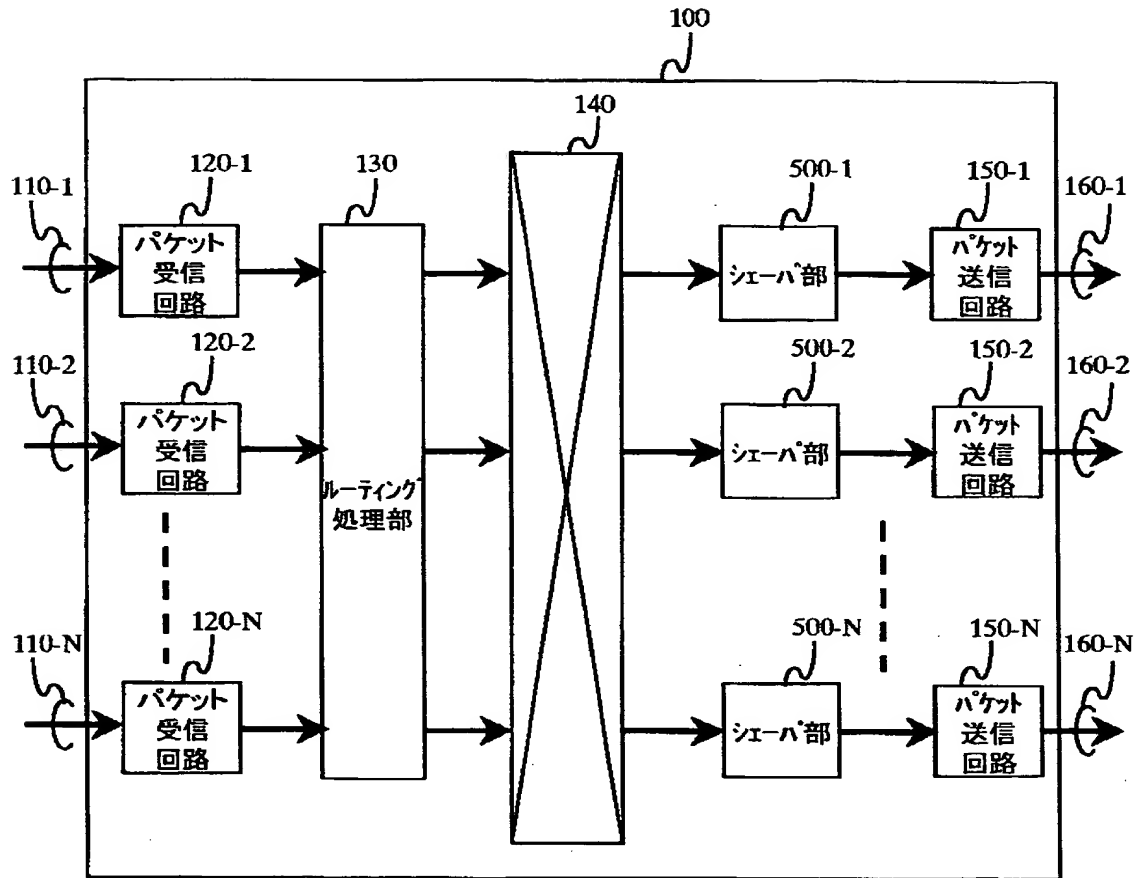
- 11…パケットヘッダ情報
- 12…キューイング優先度
- 13…廃棄判定結果
- 14…パケット送信起動信号
- 15…フロー識別子情報
- 16…パケット蓄積量情報
- 17…パケット重要度情報

18…帯域監視結果情報。

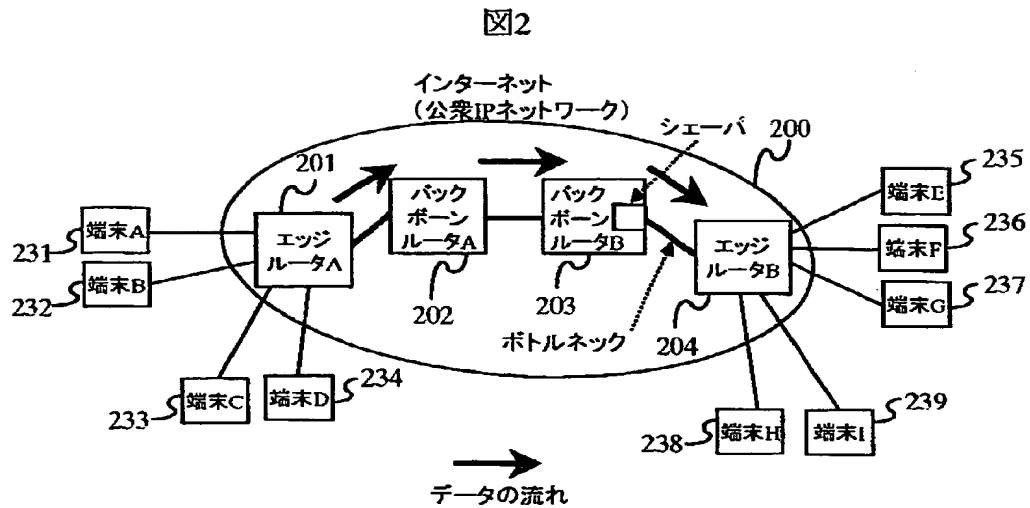
【書類名】 図面

【図 1】

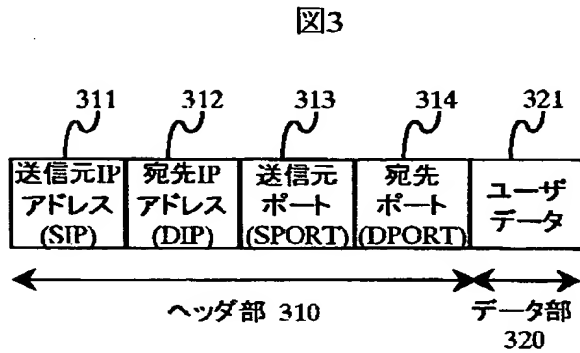
図1



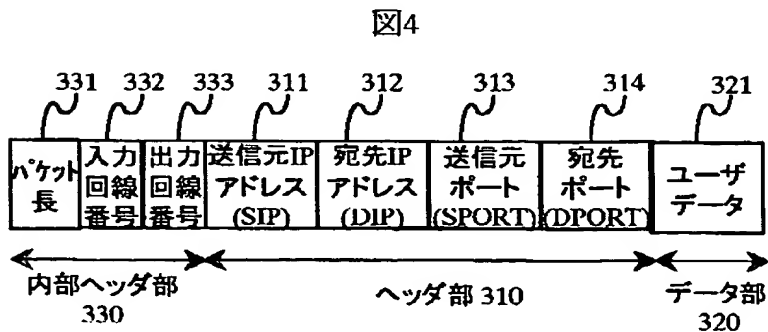
【図 2】



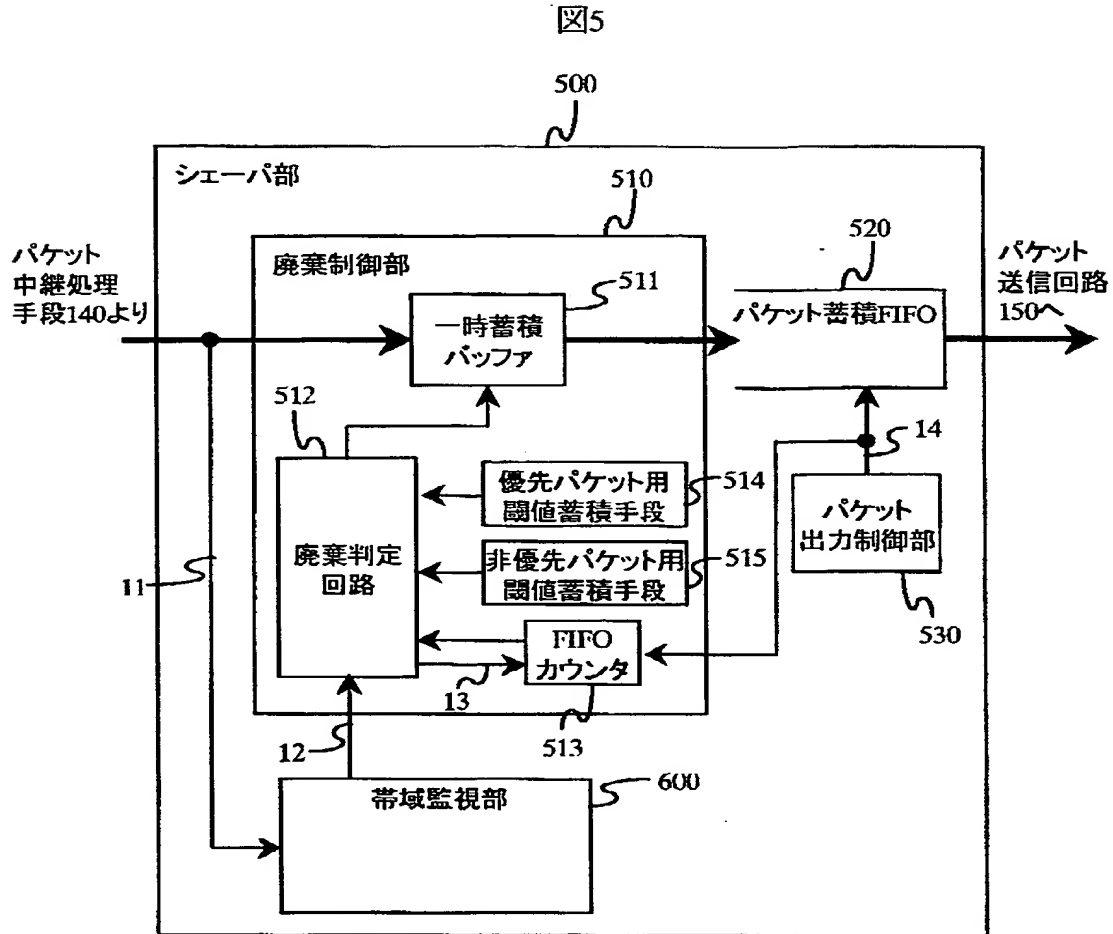
【図 3】



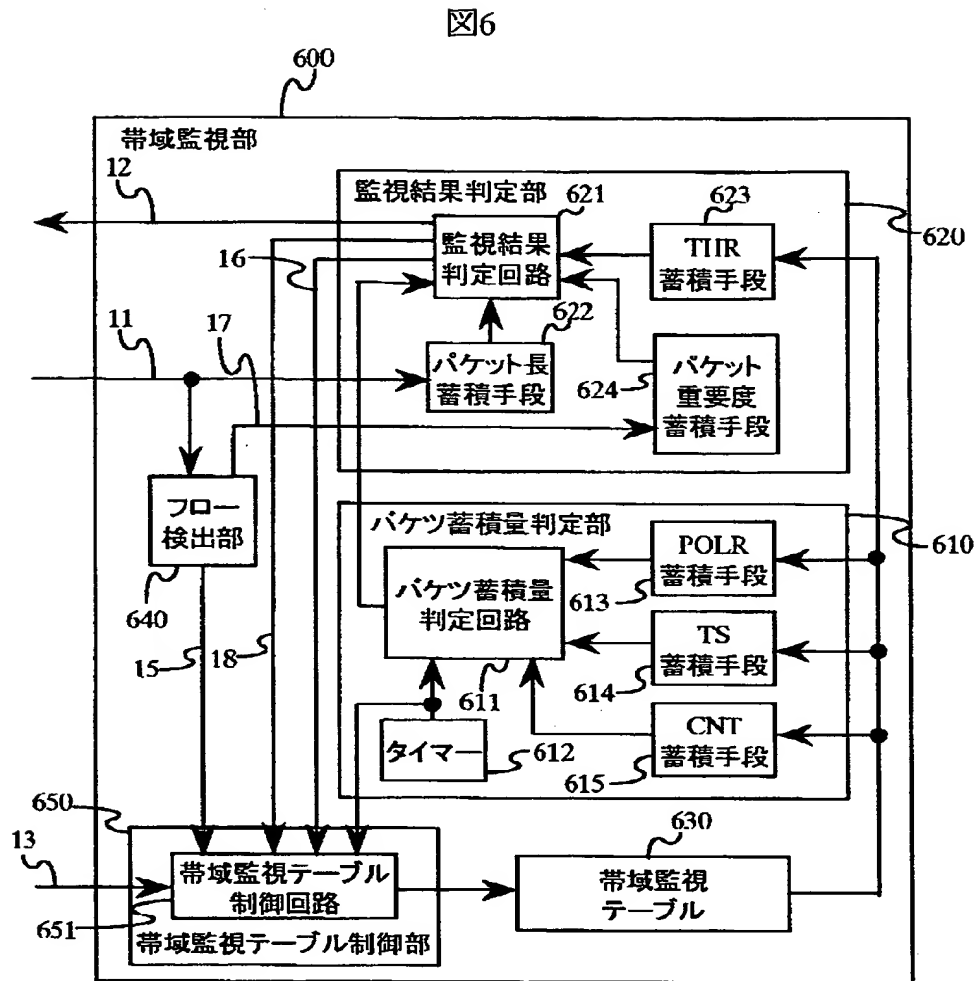
【図 4】



【図 5】

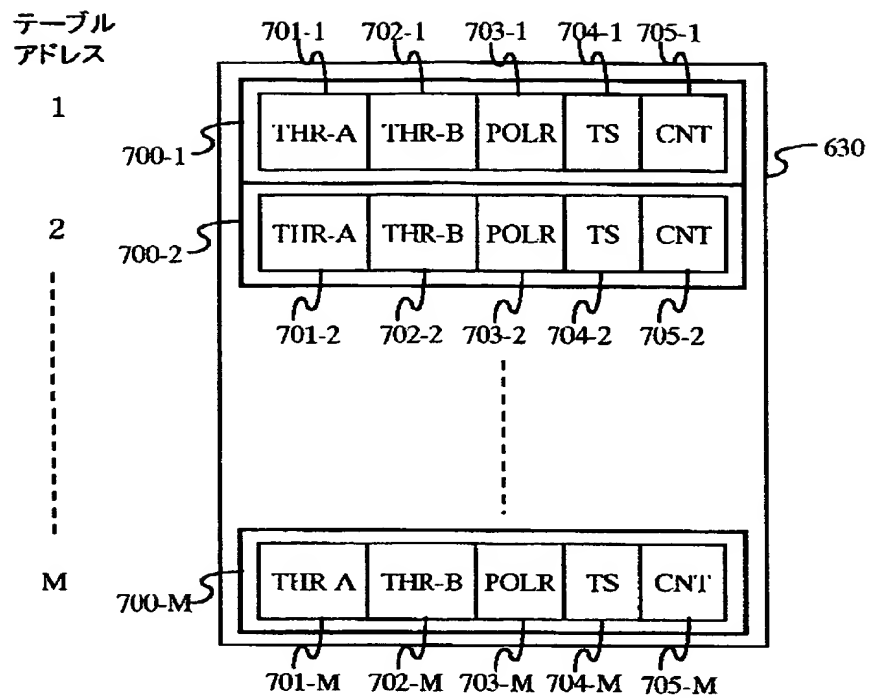


【図 6】

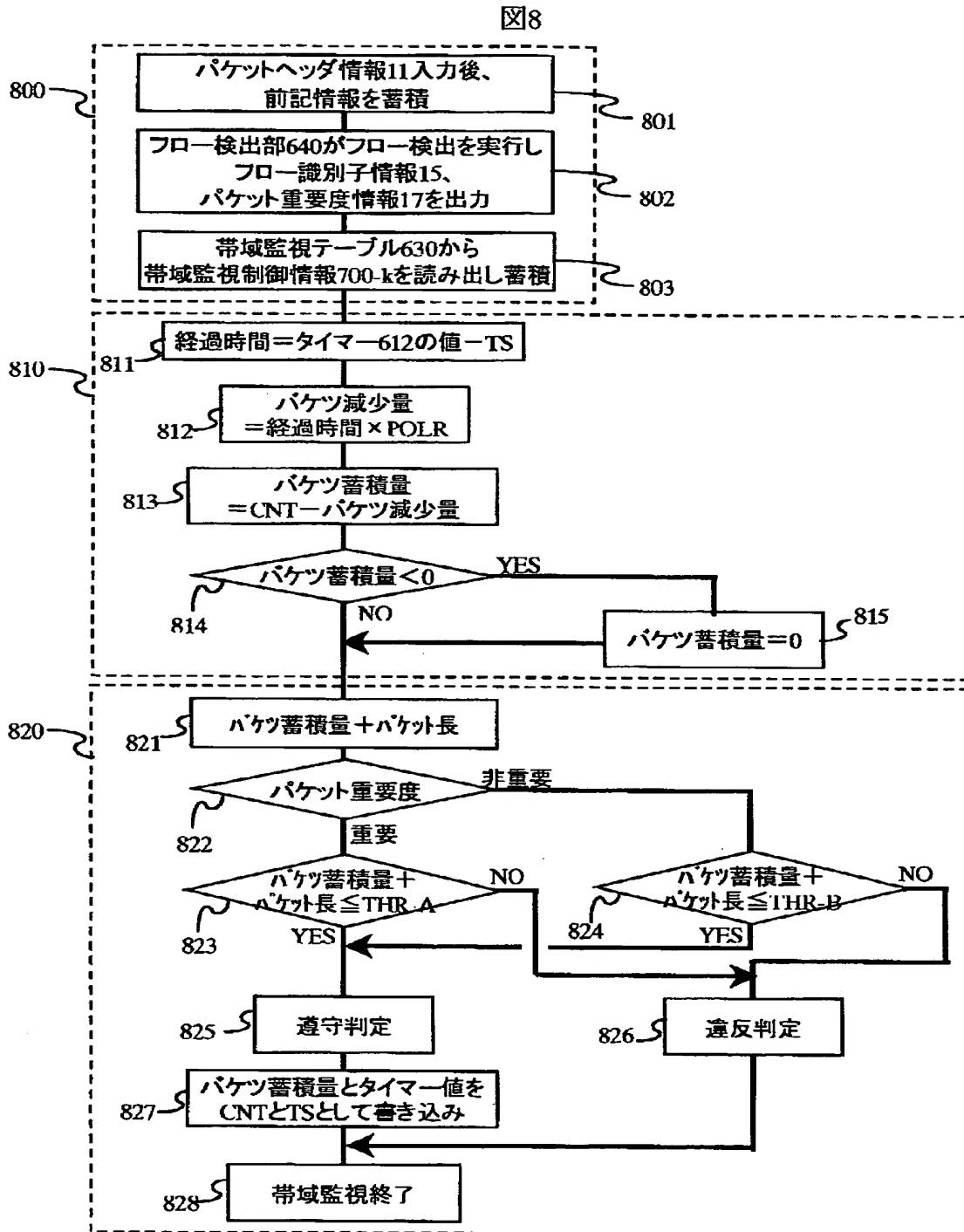


【図 7】

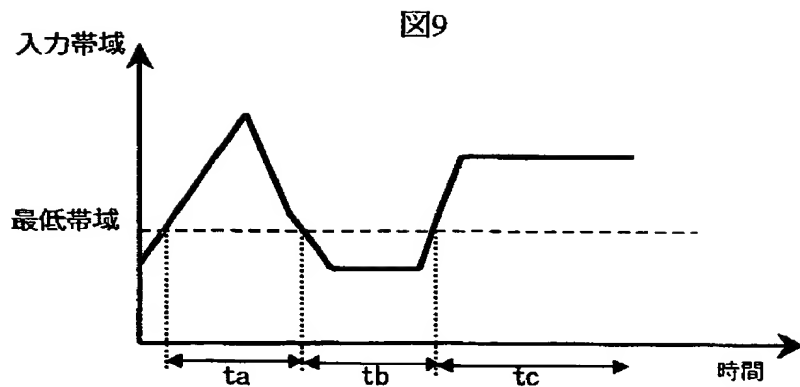
図7



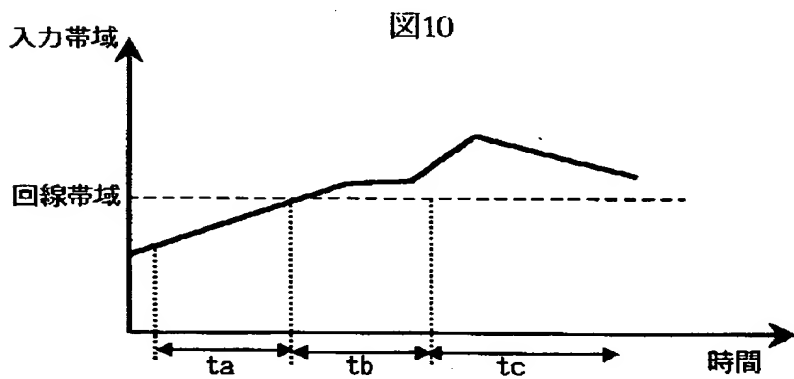
【図 8】



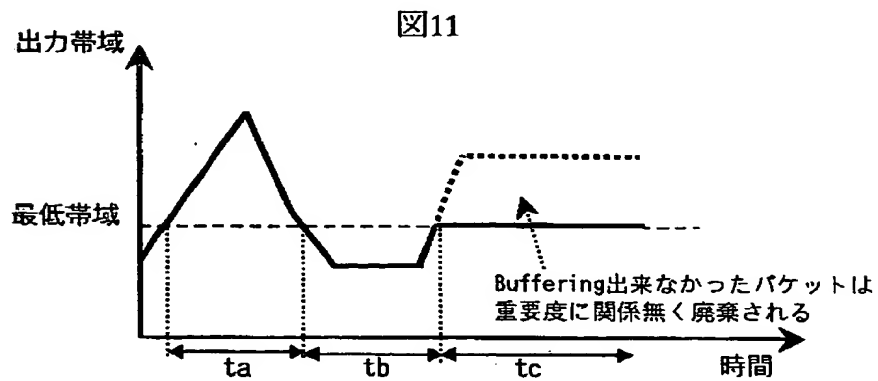
【図 9】



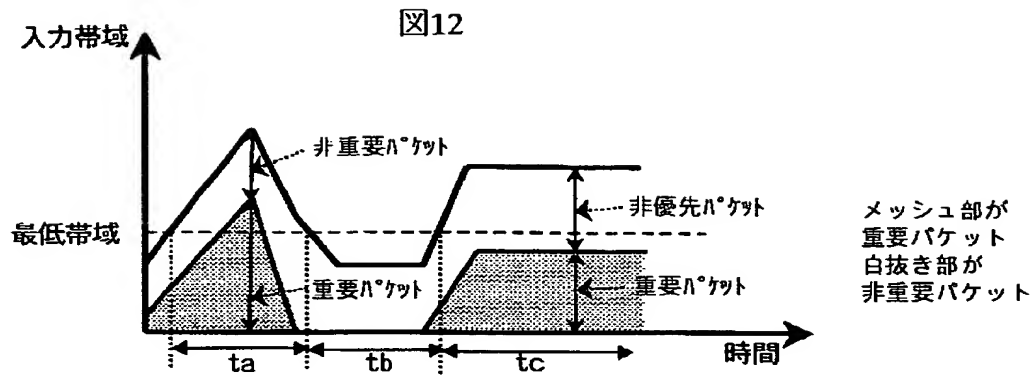
【図 10】



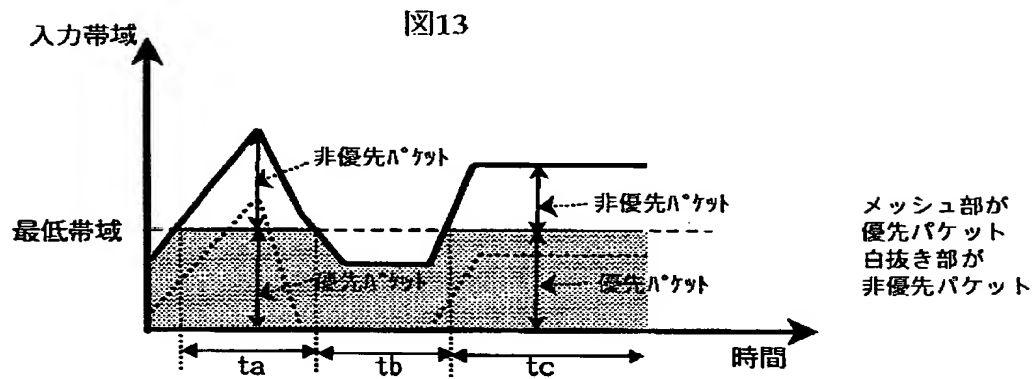
【図 11】



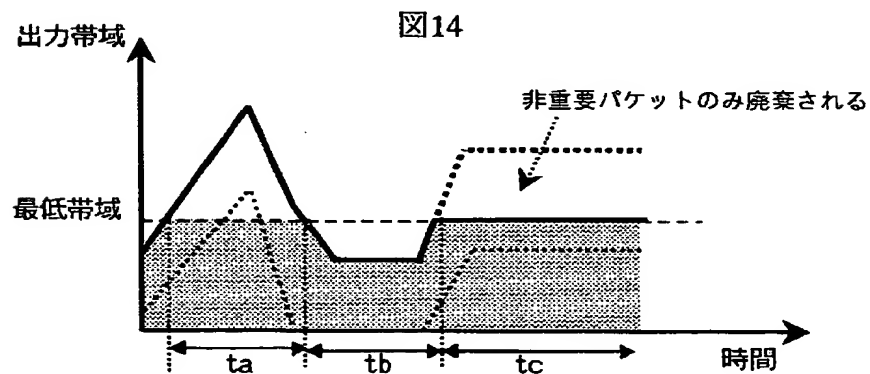
【図 12】



【図 13】



【図 14】



【書類名】 要約書

【要約】

【課題】 ユーザ毎の最低帯域を保証するシェーピングを、ユーザ数増加時にも高速実行できるシェーパを提供する。

【解決手段】 シェーパ部500にパケットが到着すると、廃棄制御部510は到着パケットの「蓄積」「廃棄」を判定する。パケット蓄積FIFO520は「蓄積」と判定したパケットを蓄積し、各ユーザの最低帯域の総和以上で送信する。帯域監視部600はユーザ毎に帯域をチェックして、最低帯域以内のパケットを優先パケットと、それ以外を非優先パケットと判定する。

パケット蓄積FIFO520に大量にパケットが蓄積すると、廃棄制御部510は非優先パケットを優先的に「廃棄」と判定して優先パケットを保護する事により、シェーパ部500は各ユーザの最低帯域を保証する。

【効果】 ユーザ毎の最低帯域を保証するシェーピングを、ユーザ数増加時にも高速実行できるシェーパを提供できる。

【選択図】 図5

出 願 人 履 歴 情 報

識別番号 [0 0 0 0 0 5 1 0 8]

1. 変更年月日 1 9 9 0 年 8 月 3 1 日

[変更理由] 新規登録

住 所 東京都千代田区神田駿河台4丁目6番地

氏 名 株式会社日立製作所